

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
2 August 2001 (02.08.2001)

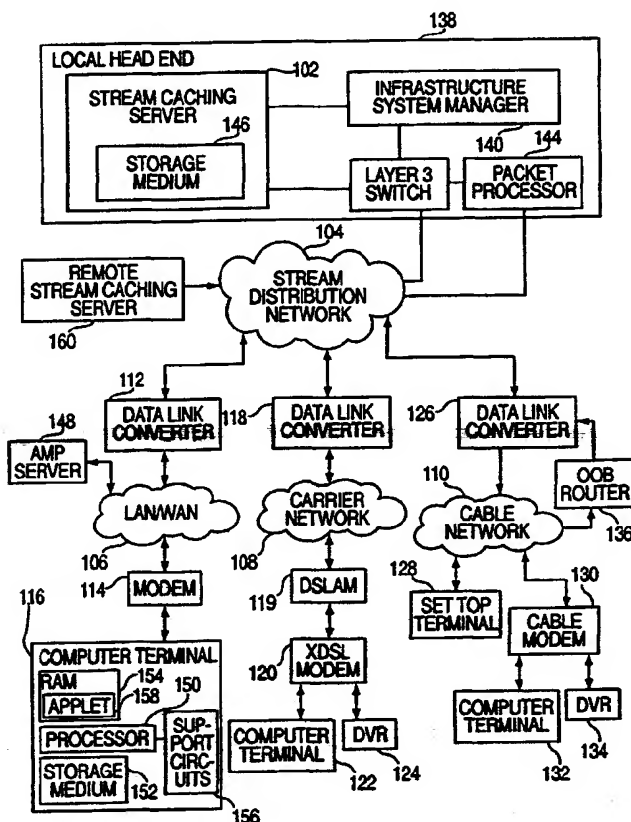
PCT

(10) International Publication Number  
**WO 01/55877 A1**

- (51) International Patent Classification<sup>7</sup>: **G06F 15/16**
- (21) International Application Number: PCT/US01/02802
- (22) International Filing Date: 29 January 2001 (29.01.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |            |                              |    |
|------------|------------------------------|----|
| 60/178,795 | 28 January 2000 (28.01.2000) | US |
| 60/178,810 | 28 January 2000 (28.01.2000) | US |
| 60/178,857 | 28 January 2000 (28.01.2000) | US |
| 60/178,809 | 28 January 2000 (28.01.2000) | US |
| 09/772,287 | 29 January 2001 (29.01.2001) | US |
| 09/772,288 | 29 January 2001 (29.01.2001) | US |
- (71) Applicant: **DIVA SYSTEMS CORPORATION**  
[US/US]; 800 Saginaw Drive, Redwood City, CA 94063 (US).
- (72) Inventors: **SON, Yong, Ho**; 535 Arastradero #310, Palo Alto, CA 94110 (US). **GOODE, Christopher, W., B.**; 722 Creek Drive, Menlo Park, CA 94025 (US).
- (74) Agents: **MOSER, Raymond, R., Jr.** et al.; 595 Shrewsbury Avenue, Suite 100, Shrewsbury, NJ 07702 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: A SYSTEM FOR PREPROCESSING CONTENT FOR STREAMING SERVER



(57) Abstract: A method and apparatus for preprocessing and postprocessing content in an interactive information distribution system. Content is retrieved from a storage medium (152) and encapsulated in accordance to an Internet Protocol (IP) format. The encapsulated content is then uploaded for storage (146) in a stream caching server (102) and for future streaming of content to different types of access networks.

WO 01/55877 A1



(84) **Designated States** (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— with international search report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**A SYSTEM FOR PREPROCESSING CONTENT FOR STREAMING SERVER**CROSS-REFERENCE TO RELATED APPLICATION

5           This invention claims benefit of U.S. Provisional Patent Applications Serial Nos. 60/178,795, 60/178,809, 60/178, 810 and 60/178,857, all filed on January 28, 2000, and such applications are herein incorporated by reference in their entireties.

          This invention is related to simultaneously filed U.S. Patent Applications Serial Nos. \_\_\_\_\_ (Attorney Docket No. DIVA 253) and \_\_\_\_\_ (Attorney  
10 Docket No. 256), filed on the same date as this application, and such applications are herein incorporated by reference in their entireties.

BACKGROUND OF THE INVENTION

15   1.     Field of the Invention

          The invention relates to electronic storage and transmission of content. More particularly, the invention relates to a method and apparatus for preprocessing and postprocessing content in an interactive information distribution system.

20   2.     Description of the Background Art

          Information systems such as video on demand (VOD) systems are capable of streaming program content to a great number of users or subscribers. To provide a requested program content to a subscriber, the VOD system retrieves the requested program content from a video server, streams the content over a stream distribution network, and  
25 converts the content to an access network that is coupled to a particular neighborhood of subscriber terminals. The user then views the requested program content at the subscriber terminal.

          However, the different types of access networks have different limitations with respect to transmission latency, bandwidth, and the like. To service a wide subscriber base,  
30 the VOD systems currently implement different solutions for each type of access network. For example, VOD systems that provide web-based video content must account for a public

and private wide area networks that support content of a particular quality of service (QoS), typically medium latency, low bandwidth and poor quality video, e.g. high jitter.

Additionally, VOD systems that provide cable-based video must account for cable networks that support low latency, high bandwidth and high quality video.

5           One example of using different solutions involves the use of separate video servers for each type of access network. Such a solution only increases the cost of providing program content at the head end. Therefore, there is a need in the art to provide a scalable VOD solution that is common for the different types of access networks. Additionally, there is a need to preprocess and postprocess content for such a VOD solution.

10

### SUMMARY OF THE INVENTION

The invention provides a method and system for preprocessing content for a stream caching server in an interactive information distribution system. In one embodiment, the method initially retrieves content in a subscriber terminal. The retrieved content is  
15 encapsulated in accordance to an Internet Protocol (IP) format used for streaming content to various access networks. The encapsulated content is then uploaded for storage in a stream caching server and for future streaming of content to different types of access networks.

The present invention preprocesses content into a common format suitable for a stream caching server capable of transmitting content to different types of access networks.  
20 In one embodiment of the invention, a user executes an applet program to preprocess content stored in a computer terminal. Such a configuration enables a user to upload content over the Internet for storage in a stream caching server and subsequent streaming to other subscribers. The invention also postprocesses content into a format supported by a particular type of player and access network used to receive the content from the stream  
25 caching server.

### BRIEF DESCRIPTION OF THE DRAWINGS

The teachings of the present invention can be readily understood by considering the following detailed description in conjunction with the accompanying drawings, in which:

FIG. 1 depicts a high level block diagram of a first portion of an interactive information distribution system embodied in the present invention;

FIG. 2A depicts one embodiment of an Internet Protocol (IP) packet used in the information distribution system of FIG. 1;

5        FIG. 2B depicts one embodiment of a Realtime Transport Packet (RTP) contained in a payload section of the IP packet of FIG. 2A;

FIG. 3 depicts a data structure useful in understanding an embodiment of the present invention;

FIG. 4 depicts a flow diagram of a method for implementing the preprocessing of  
10        program content in accordance to one embodiment of the present invention; and

FIG. 5 depicts a flow diagram of a method for and postprocessing content in accordance to another embodiment of the present invention.

To facilitate understanding, identical reference numerals have been used, where possible, to designate identical elements that are common to the figures.

15

#### DETAILED DESCRIPTION

FIG. 1 depicts a high level block diagram of an interactive information distribution system 100. One application of the distribution system 100 is as a video of demand (VOD) system, as described in U.S. Patent Application No. 08/984,710, filed December 3, 1997  
20        and incorporated herein by reference. In such a VOD system 100, a user may request and receive a particular content selection, e.g., video, movie or programming content, from a service provider without any time restrictions associated with normal cable programming.

The information distribution system 100 comprises a stream caching server 102, a stream distribution network 104, at least one access network and at least one subscriber  
25        terminal. The stream caching server 102 receives, stores and streams content in accordance to an Internet Protocol (IP). One example of such a stream caching server 102 is disclosed in simultaneously filed U.S. Application \_\_\_\_\_, attorney docket DIVA 253, entitled "Method and Apparatus for Streaming Content in an Interactive Information Distribution System, which is herein incorporated by reference. The content is configured within a  
30        payload portion of each IP packet received, stored and streamed by the stream caching

server 102. The use of this IP formatted content enables a single stream caching server 102 to stream content via an integrated stream distribution network 104 to different types of access networks. As such, the system 100 is capable of streaming the same content to any cable service subscriber or any person using the Internet.

5 In accordance to the present invention, the stream caching server 102 may receive content that is preprocessed at, for example, a remotely located subscriber terminal. One such subscriber terminal is a computer terminal 116 comprising a processor 150, a storage medium 152, a random access memory (RAM) 154 and support circuits 156. The RAM 154 stores an applet 154 that is downloaded from, for example, a HTTP server 148 coupled  
10 to an access network, for example., a private local area network (LAN) or wide area network (WAN) 106. The processor 150 executes the applet 154 to initiate the preprocessing in the present invention. The storage medium 152 stores the content to be preprocessed. The support circuits 156 provide an interface for receiving the applet 154 from the http server 148, receiving content from the streaming cache server 102, or  
15 uploading preprocessed content to the http server 148.

Possible configurations of the applet 154 include a JAVA Applet Plug In for an Internet browser, or a software program written in a particular programming language, e.g., C++. Once the processor 150 executes the applet 154, the content is retrieved from the storage medium 152. The content may include standard multimedia files in a variety of  
20 formats, e.g., AVI (Audio Video Interleaved), Moving JPEG, MPEG-1, MPEG-2, MPEG-4, MP3, Quicktime, and the like. If a need exists to convert the content into a particular format, the retrieved content may be transcoded into a format supported by a viewer or subscriber terminal that eventually receives the downstream content. The transcoding of content changes the format and rate of the retrieved content. One example of such  
25 transcoding is the conversion of MPEG-2 content into MPEG-4 content that can be played on a graphic processor in set top terminals or personal computer (PC) terminals. Other types of content may be transcoded into MPEG-2 content playable on conventional set top terminals. Some transcoding requires decoding to baseband followed by encoding according to the desired format. Some transcoding may be performed without baseband  
30 decoding.

The content, whether transcoded or not, is then encapsulated into a format that is optimal for the stream caching server 102. In one embodiment, the content, as MPEG-2 packets, is encapsulated in a payload of each Realtime Transfer Protocol (RTP) packet contained in the payload of an IP packet. The format of such encapsulated content is shown in FIGS. 2A-2B. However, the use of the RTP packet also supports non real time applications.

FIG. 2A depicts one embodiment of an Internet Protocol (IP) packet 200 used in the present invention. The IP packet comprises an IP header 210 and an IP payload 320. The IP payload comprises a UDP (User Datagram Protocol) packet 221 comprising a UDP header 222 and a UDP payload 224. A RTP packet 230, a stream integrity check 226 and a cyclic redundancy check (CRC) 228 is illustratively contained in the UDP payload 224. In one embodiment of the IP packet 200, the IP header 210 is 20 bytes, the UDP header 222 is 8 bytes, the stream integrity check field 226 is 4 bytes and the CRC field 228 is 4 bytes.

FIG. 2B depicts one embodiment of a Realtime Transport Packet (RTP) 230 contained in a payload section 220 of the IP packet 300 of FIG.2A. The RTP packet 230 comprises a RTP header 240 and a RTP payload 250. Five MPEG-2 packets 352, 354, 356, 358 and 360 are illustratively contained in each RTP payload 250. The number of MPEG-2 packets in the RPT payload 250 corresponds to the buffer space in the Fibre Channel controller in the packet processor 144.

For the embodiment shown in FIGS. 3A-3B, the transcoding may also include compression of the IP header 310, the UDP header 322 and the RTP header 340. The compression of these headers 310, 322 and 340 optimizes the storage of content on the storage medium 146 of the stream caching server 102. Additionally, the transcoding may include encryption of the content.

Compression of the IP header 310 may include compression of non-address fields, and deletion of source and destination IP addresses. The IP source address is subsequently decompressed (at the packet processor 144) as the IP address at the output interface of the stream caching server 102. The IP destination address is assigned prior to streaming and is based upon the data link converter 112, 118 or 126 or edge device for the target access network 106, 108 or 110.

The compression of the UDP header 322 may delete source and destination port numbers in the storage medium 146. New values are then assigned to the source and destination port numbers prior to streaming of content over the stream distribution network 104. The source port number is assigned a unique stream number with IP addresses. The  
5 destination port number is assigned a unique target stream number within the data link converter 112, 118 or 126 or the target edge device.

Once encapsulated into the IP format, the content is then uploaded to the http server 148 via the modem 114 and the LAN/WAN 106. The HTTP server 148 provides a user interface, e.g., a HTML (HyperText Markup Language) page, for the user to upload the  
10 content to the stream caching server 102. The http server 148 also transmits the transcoded content upstream to the data link converter 112 via the LAN/WAN 106. The data link converter 112 modulates the encapsulated content for upstream transmission over the distribution network 104 for storage in the stream caching server 102.

The preprocessing of the present invention is not limited to a computer terminal 116  
15 uploading content over a LAN/WAN 106. For example, the encapsulation may also occur in the computer terminal 116 prior to uploading to content to the http server 148. Additionally, the preprocessing may be initiated from a computer terminal 122 or digital video recorder 124 over a carrier network 108, e.g., a T-1 or T-3 line. As such, a user of any computer terminal 116 and 122 author multimedia content over a network, e.g.,  
20 Internet, and store the content in a virtual video shelf at the stream caching server 102 for playback by other users. The preprocessing is also applicable to content from a content provider such as a movie manufacturer.

The preprocessing may also include the creation of metadata once the processor 150 executes the applet 154. The metadata generally contains a variety of information about the  
25 content to be stored on the stream caching server 102 and streamed to a viewer or subscriber terminal. One embodiment of the metadata is in the form of a data structure that is prepended prior to a file associated with the content.

The metadata may comprise many different types of information used by the stream caching server 102 in steaming the content to a viewer device or subscriber terminal. One  
30 type of metadata that identify the content include title, author, screenwriter, actors, length



of play, timing information, play rate, e.g., constant bit rate (CBR) or variable bit rate (VBR), genre of content, size of content, and the like. Other forms of metadata indicate the type of content and the type of player used to play the content. Illustrative types of content include AVI, MPEG-1, MPEG-2, MPEG-4, MP3, Quicktime, and Moving JPEG. The  
5 metadata may also include MPEG7 structure including scene descriptions and indices. Exemplary types of player devices include MPEG-1 player, MPEG-2 player, MPEG-4 player, Microsoft Media Player, Real Video / Real Audio Player, and QuickTime Player.

The metadata may include pricing information and restrictions to view the content from the server 102. A user or service provider may preset the price and access restrictions  
10 that are required to view the preprocessed content from the server 102. Pricing information include a price to view the content, applicable discounts associated with viewing the content, and applicable package deals when ordering the preprocessed content with other content selections. Access restrictions include rating of content, viewing window information and sales window information. The viewing window is a graphical interface  
15 that requires a subscriber to enter a correct password to receive the content from the server 102. The sales window is a graphical interface that requires a subscriber to pay for viewing a particular content selection.

The metadata comprises indexing at IP packet boundaries, e.g., Group of Pictures (GOP) boundaries and frame boundaries. This indexing enables the stream caching server  
20 102 in responding to an interactive VCR like commands, e.g., fast forward (FF), rewind (REW), pause, stop, bookmark, and return to place. Specifically, the indexing information supports random frame access at a content stream granularity, separate FF and REW tracks, random frame access of FF and REW tracks, pause/play, bookmarks for each active subscriber, and DVD scene selection. Additionally, the metadata may also include markers  
25 for changes in variable bit rate (VBR) and statistical multiplexing.

The indexing enables the use of MPEG-7 based descriptors for indexed access of content. The indexing of content can be on a per frame basis or a per GOP basis. MPEG-7 based descriptors include a length of descriptor, i.e., from a start frame to an end frame, and a schema for a database of indices. The database is a hierarchical based database that

allows for hierarchical scene description. For example, a root index may represent a scene of Paris, while a branch index may represent a scene of the Eiffel Tower.

The indexing is also used for stream creation. In one embodiment, the stream is created in realtime from a single MPEG-2 or MPEG-4 content stream, e.g., a start GOP to an end GOP, or a start frame to an end frame. In a second embodiment, the stream is created in non real time and the modified stream file is stored. In a third embodiment, the stream is transcoded or re-encoded such that a reference frame or I-frame is forced to be the frame with information desired in an index, even if the frame arrived in the packet processor 144 as a predictive frame, e.g., B-frame or P-frame. Stream indexing will also be discussed below with respect to FIG. 3.

Once the preprocessed content is uploaded, the stream caching server 102 may store and stream the preprocessed content, or alternatively stream the content in real time. The streaming of content encapsulated in IP format enables the stream caching server 102 to stream content to subscribers via different types of access networks 106, 108 and 110. As such, only one stream caching server 102 and one distribution network 104 is required to provide scalable streaming. Namely, one stream caching server 102 may stream content to the LAN/WAN 106, the carrier network 108, the cable network 110 and any other access network that supports IP. This greatly reduces the hardware cost at the head end 138, as the prior art requires a streaming caching server 102 and a distribution network 104 for each type of access network 106, 108 and 110.

For example, the server 102 may stream content to a private Local Area Network (LAN) or Wide Area Network (WAN) 106. Specifically, the system 100 may stream content through the stream distribution network 104 to a digital link converter 112, the LAN or WAN 106, a modem 114 and a display device coupled to a computer terminal 116. If a user decides to request content from the server 102 or uplink content, e.g., a home movie, to the server 102, that request or content would travel upstream in a path reverse to that of the downstreamed content. The digital link converter 112 modulates the program content for transmission via the private LAN/WAN 106. Additionally, the data link converter 112 extracts a MPEG formatted program content from the RTP formatted stream from the stream distribution network 104 and transcodes the program content into a format

that is supported by the LAN/WAN 106. One example of the data link converter 112 is a DIVA Digital Link (DDL 500) that performs quadrature amplitude modulation (QAM) on the downstream program content. The LAN or WAN 106 is a private network provided by a private party or an Internet Service Provider (ISP). The modem 114 demodulates video  
5 content for viewing on the computer terminal 116.

The server 102 may also stream content via the stream distribution network 104 to a data link converter 118, a carrier network 108, a digital subscriber line access multiplexer (DSLAM) 119, an x-DSL modem 120 to either a computer terminal 122 or a digital video recorder (DVR) 124. A request for content or upload of content would travel in the reverse  
10 path taken by the downstream content. The carrier network 108 may include a T-1 or T-3 transmission link. The data link converter 118 multiplexes the downstream content for transmission via the carrier network 108. Additionally, the data link converter 118 may extract MPEG packets from the IP formatted stream from the stream distribution network 104. The DSLAM 119 demultiplexes the downstream content to a particular xDSL modem  
15 120. The xDSL modem 120 demodulates the content for viewing on a computer terminal 122 or a display device (not shown) coupled to the DVR 124. The xDSL modem 120 may comprise a ADSL (asynchronous digital subscriber line) modem, a VDSL (very high data rate digital subscriber line), and the like.

The server 100 may also stream content (or program content selection) via the  
20 stream distribution network 104 to a data link converter 126, the cable network 110 to either a set top terminal 128 or a cable modem that is coupled to a computer terminal 130 or a DVR 132. The data link converter 126 operates in a similar manner to the digital link converter 112 except to format the content for transmission in the cable network 110. The content is transmitted from the cable network 110 to a set top terminal 128 or a cable  
25 modem 130 that demodulates the program content for viewing on a computer terminal 132 or a display device coupled to the DVR 134. A request from a cable subscriber or user is processed via the cable network 110, the OOB (out of band) router 136 and the modulator 126 that modulates the request back to the stream distribution network 104.

Although the system 100 is illustratively shown to stream program content to the  
30 LAN/WAN 106, the PSTN 108 and the cable network 110, the system 100 may also stream

content to other types of access networks. For example, the system 100 may also stream program content to satellite and terrestrial networks. Additionally, each system 100 actually streams content over many more access networks and subscriber terminals than the example shown in FIG. 1.

5           The stream caching server 102 is located at the local head end 138 with an infrastructure system manager 140, a switch 142 and a packet processor 144. The stream caching server 102 comprises a storage medium 146 to store the content preprocessed in accordance to the present invention. One configuration of the storage medium 146 is a redundant set of disk arrays, e.g., Redundant Array of Inexpensive Disks (RAID).

10           The infrastructure system manager 140 coordinates a (user) request from the subscriber terminal by passing the request to the stream caching server 102 and establishing a session between the subscriber terminal and the stream caching server 102. An exemplary infrastructure system manager 140 is the DIVA System Manager (DSM). As disclosed in U.S. Application \_\_\_\_\_, Attorney docket DIVA 256, entitled  
15           “Method and Apparatus for Managing an Integrated Information Distribution System”, which is fully incorporated by reference in its entirety. The switch 142 routes the user request from the stream distribution network 104 to the system manager 140. Additionally, the switch 142 routes the retrieved content from the stream caching server 102 to the packet processor 144.

20           The storage medium 148 stores the preprocessed content in an IP format. The content is configured as a plurality of MPEG, e.g., MPEG-2 or MPEG-4, packets contained in a payload of a RTP packet within an IP packet. For example, the payload of each RTP packet may contain five MPEG-2 packets. The structure of the IP packet is shown to FIG. 3B. The RTP format (RFC 1889) minimizes the latency in streaming content from the  
25           server, by supporting the streaming of content in real time. Additionally, the content in the IP packet can be configured to have a minimal Quality of Service (QoS), e.g., data latency.

          The packet processor 144 postprocesses the content into a format supported by a particular type of player and access network 106, 108 and 110 used to receive the content from the stream caching server 102. Such a player is either a software module downloaded  
30           from a HTTP server 116 to a computer terminal 122, a hardware module coupled to a

subscriber terminal, or a card inserted into a subscriber terminal. Exemplary players include a MPEG-1 player, a MPEG-2 player, a MPEG-4 player, a Microsoft Media Player, a Real Video/Real Audio Player, a QuickTime Player, a Wireless Device Video or Audio Player, and the like.

5           The packet processor 144 transcodes the content is performed without disturbing the IP format. For example, the packet processor 144 separates the content, e.g., MPEG-2 packets, and header information in the IP packet, transcodes the content packets into a desired format supported by the access network and downstream player, and combines the transcoded packets with the header information to recreate the IP packet. Such transcoding  
10 is performed at an elementary packet level for transmitting at the transport packet level. Additional functions performed by the packet processor 144 include jitter correction, creating of a PES (packet elementary stream), stream splicing, and statistical multiplexing.

          More specifically, the transcoding includes the conversion content in the RTP payload into a format suitable for the access network 106, 108 and 110, but the transcoded  
15 content is still encapsulated in the IP packet stream. Such transcoding may change the format and rate of the content. For example, the transcoding may include the conversion of MPEG-2 formatted content into MPEG-1, MPEG-4, AVI, Moving JPEG, MP3, Quicktime, Wireless Applications Protocol content, and the like. The transcoding is performed in accordance to an extended Real Time Streaming Protocol (RTSP - RFC 2326) such that  
20 stream manipulations conform to Internet standards and are applicable to any access networks that support IP.

          Additionally, the exact manner of the transcoding depends on the available bandwidth in the access network used to receive the content at the player. For example, the packet processor 144 may perform statistical multiplexing to dynamically allocate the  
25 amount of available bandwidth for streaming content to a particular viewer. To perform such statistical multiplexing, the packet processor 144 may stream content at either a constant bit rates or variable bit rates.

          The transcoding is also adjustable to bandwidth degradations. To process lossy video, the transcoding may include lossy filtering within frames of content, dropping  
30 frames of content, e.g, resulting in a playback rate of 30 frames per second to 15 frames per

second, and delivering still frames that contain important information. For non-lossy compression, the transcoding may include dropping MPEG null packets, and transcoding or re-encoding content to an acceptable quality.

The packet processor 144 may automatically perform such transcoding, or perform  
5 transcoding in accordance to user configured preferences. These preferences may include choices for a particular player, e.g., formatting, play rates, and type of conversion or transcoding. For example, if the player is embodied in software in a PC, then the content is transcoded into MPEG-2 format. However, if the player is a hand held device, then the content is transcoded into JPEG or MPEG-4 format. Additionally, the transcoding may be  
10 dynamically performed based on a user preference profile. Such a profile is based either on history or a default preference. For example, if the player is in the PC, the packet processor 144 transcodes the content into MPEG-2 at 4 Mbps and a constant bit rate.

The content in the payload of each RTP packet is sized to minimize the latencies in streaming content from the stream caching server 102 to the distribution network 104. The  
15 read block for the packet processor 144 is sized to the MPEG packets in the payload of each RTP packet. The number of MPEG packets in each RTP packet is constrained by an available buffer space in a Fibre Channel controller that is used to read the content.

The content streamed by the stream caching server 102 is not limited to content previously stored in the storage medium 148. In one embodiment, the stream caching  
20 server 102 streams content from another remotely located server, i.e., a server located at a remote headend. Such a configuration is further described with respect to FIG. 2.

The manager 140 provides session management for streaming content in accordance to the RTP Control Protocol (RTCP). Such management is particularly important in the case of content streamed to the local stream caching server 102 from the remote server. If  
25 any errors occurred during the streaming from the remote server, these errors are multiplied when the cached or stored content is then streamed to the many subscribers. RTCP enables the detection and transmission of only the read blocks affected by the streaming errors.

FIG. 2 depicts another portion of the interactive information distribution system 100 of FIG. 1. This portion of the system 100 comprises the stream caching server 102 and the  
30 infrastructure system manager 140 at the local head end 138, a stream caching server 202

and an infrastructure system manager 204 at a remote head end 206, and a backbone streaming network 210. The stream caching server 202 and the infrastructure system manager 204 at the remote head end 206 operates in a similar manner to the respective stream caching server 102 and the infrastructure system manager 140 at the local head end 5 138 that were previously described.

The local infrastructure system manager 140 receives a request for a particular content selection and determines whether a user requested content selection is stored in the storage medium 148. If the request content is not in the storage medium 148, the local infrastructure system manager 140 identifies a remote stream caching server 202 that stores 10 the requested program content and provides a (server) request to the remote system manager 204. For example, a local system manager 140 in San Francisco may request content from another remote remotely located server 202 in Boston.

In response to this server request, the local system manager 140 coordinates the streaming remote stream caching server 202 streams the requested program content over 15 the backbone streaming network 210 to the local stream caching server 102. The content is then streamed to the subscriber. If the local system manager 140 determines that there are enough user requests above some predetermined threshold number, then the content from the remote stream caching server 202 is also stored in the local stream caching server 102.

FIG. 4 depicts a flow diagram of a method 400 for implementing the preprocessing 20 of content e.g., a video program, in accordance to one embodiment of the present invention. The method 400 assumes that a user or subscriber has already downloaded an applet 154 from a http server 115. Specifically, the method 400 starts at step 402 and proceeds to step 404 where preprocessing is initiated when the applet 154 is executed by the processor 150 in the computer terminal 116. At step 406, a query determines whether a user has 25 purchased shelf space. Namely, step 406 determines whether the user has purchased storage space or use of a portion of the storage medium 146 at the stream caching server 102. Step 406 may also cover situations where a user has access to shelf space.

If the user has not purchased shelf space on the storage medium 148, the method 400 proceeds to end at step 424. If the user has already purchased shelf space on the 30 storage medium 148, the method 400 proceeds to step 408, where content is loaded from a

memory 152 of the computer terminal 116. The content may comprise any multimedia presentation, e.g., a home movie, created by the user.

At step 410, the method creates metadata for the loaded content. The metadata contains indexing information that enables the stream caching server 102 to respond to user  
5 interactivity commands within a group of pictures or approximately one half a second. Examples of user interactivity commands include fast forward (FF), rewind (REW), pause, stop, bookmark and return to place. Additionally, the metadata includes information that enables the stream caching server 102 to determine the type of file and the resolution of the content.

10 The method 400 proceeds to step 412 where a query determines whether to transcode content. If no transcoding is required, the method 400 proceeds to step 416.

If transcoding is required, the method 400 proceeds to step 414 where the content is transcoded into a format that is supported by a viewer used to view downstream content. Step 414 may convert both the format and bit rate of the program content. In one  
15 embodiment, the content may be transcoded (at a elementary stream level) from MPEG-1, MP3, AVI, QuickTime or Moving formed into MPEG-2 formatted packets. At step 416, the method 400 encapsulates the transcoded content, e.g., MPEG-2 format, into an IP packet format (at the transport stream level). As previously described with respect to FIGS. 2A and 2B, storage of content in this IP packet format minimizes the retrieval time when  
20 the stored content is retrieved from the storage medium 146 and stream to the distribution network 104.

The method 400 proceeds to step 418, where the transcoded content and metadata is uploaded from the computer terminal 116 and into the storage medium 148 of the stream  
caching server 102. Once the transcoded content is stored at the stream caching server 102,  
25 the method 400 enables the user (sender of the content) to establish or set permissions at step 420. This step establishes a subset of users who may access the content from the stream caching server 102. The method 400 proceeds to step 422, where a HTML page is established at the http server 115 for access by other users. The method 400 ends a step 424.



FIGs. 5A and 5B depict a flow diagram of a method for accessing program content that was preprocessed and stored at the stream caching server 102. In one embodiment of the method 500, a user may access the full version of the program content only after correctly entering the password and paying to view the content.

- 5           The method 500 starts at step 502 and proceeds to step 504, where a user access a html page and selects a particular program content to be played from the stream caching server. At step 506, the method 500 queries whether the user has entered the correct password as previously established by the owner of the program content. Step 506 may alternatively query for a particular subscriber name or keyword.
- 10           If the correct password is not entered, the method 500 ends at step 540. If the correct password is entered, the method 500 proceeds to step 508, where a query determines whether the user has a viewer or player to view the selected program content. Namely, step 508 determines whether a correct type of viewer or player is detected at the subscriber terminal, e.g., a computer terminal 116, 122 or 132. If the correct player is not
- 15           detected, the method 500 proceeds to download the player at 510 and then proceeds to step 512. If the correct player is detected, the method 500 proceeds directly to step 512.
- At step 512, a query determines whether the downloaded player supports playback (of content) at full quality. If the player does not support playing of content at full quality, the method 500 proceeds to step 526. If the player supports playing of content at full
- 20           quality, the method 400 proceeds to step 514, where a query determines whether the user has paid to view a full quality version of the content, e.g., a program content file . If the user has paid to view the full quality version of the program content, the method 500 proceeds to retrieve the full IP formatted content from the stream caching server 102 at step 516 and streams the retrieved content over the distribution network 104 at step 518. The
- 25           method 500 proceeds to step 520, where the data link converter 112, 118 or 126 extracts the content (at the transport level) from the IP packets, content or MPEG-2 packets from the IP formatted content. At step 522, a query determines whether to transcode the extracted content. Such transcoding is required to satisfy constraints in (downstream) transmitting the content over the access network 106, 108 and 110, and for playing the content on the
- 30           viewer. If transcoding is not required, the method 400 proceeds to step 536. If transcoding

is required, the method 400 transcodes the content at normal quality at step 524 and proceeds to step 536.

Referring back at step 514, if the user has not fully paid to view the full quality version of the content, the method 500 proceeds to step 526. At this step 526, the method 500 proceeds to retrieve a sample or predefined portion of the IP formatted content (file) from the stream caching server 102 at step 526. The method 500 then proceeds to stream the retrieved portion over the distribution network 104 at step 528 and extract the content, e.g., MPEG-2 packets, from the IP formatted content at step 530. The method 500 proceeds to step 532, where a query determines whether to transcode the extract content for transmission over the access network 106, 108 or 110 and for playback on the viewer or player. If no transcoding is required, the method 500 proceeds to step 536. If transcoding is required, the method 550 proceeds to transcode the partial content at low quality at step 534 and then to step 536.

At step 536, the method 500 sends the transcoded content the subscriber terminal, e.g. a computer terminal 116, 122 and 132 via the access network 106, 108 and 110. The method 500 proceeds to play either the full or partial quality content at step 538 and ends at step 540.

FIG. 3 depicts a data structure useful in understanding an embodiment of the present invention. Specifically, FIG. 3 depicts a content stream 310 including a plurality of index points denoted as  $I_1$ ,  $I_2$  and so on up to  $I_N$  (collectively index points  $I$ ). It will be appreciated that while the index points  $I_1$  through  $I_N$  are depicted as being spaced in an approximately even manner, there is absolutely no requirement for such equal spacing of the indices. Each index point represents an appropriate stream entry point beginning with, for example, an I-frame in the case of an MPEG content stream. Each indexed portion may be long or short, may comprise for example an entire scene or a plurality of scenes, other index divisions may be used.

FIG. 3 also depicts a sub-stream 320 comprising a plurality of index points denoted as  $i_1$ ,  $i_2$  and so on up to  $i_8$ . For purposes of this discussion, it is assumed that the sub-stream comprises portions of content between index point  $I_2$  and index point  $I_4$ , yet not portions of the content traversing those index boundaries. The shaded portion of content stream 310

depicts the content within the sub-stream. The shaded portions of the sub-stream 320 depict content that is not included within the sub-stream. It is noted that the sub-stream may be stored as a separate entity along with the content stream 310. The sub-stream may comprise those image frames associated with a portion of content indexed according to the MPEG-7 standards. For example, in the case of a movie having a plurality of scenes, where objects within the scenes of the movie have been indexed according to the MPEG-7 format, a particular object (e.g., a sunset, classic automobile or actor) may be associated with an object or a group of objects represented within the content stream 310. Thus, the sub-stream 320 stores content specifically associated with a desired object so that such content may be retrieved.

It is noted that in retrieving a sub-stream it is desirable for the first frame, for example an MPEG-2 stream, to comprise an intra-coded frame (an I-frame). As such, since it is possible for the desired content to be included in the content stream 310 at a point comprising a non I-frame, the sub-stream 320, when created, includes transcoding of at least the first frame within the sub-stream into an intra-frame coding format. Additionally, if desired, each of the index points within the sub-stream may be reencoded to insure that these points also comprise I-frames. It is noted that many sub-streams may be generated and stored and associated with the main content stream.

In one embodiment of the invention, local content is loaded onto a client, such as a set-top terminal or computer, from a local content source such as a video cassette recorder (VCR), digital video recorder (DVR), personal video recorder (PVR), computer or other storage and/or content source. For example, local content may be loaded onto the set top terminal or streamed to the set top terminal from a camera, live video and/or live audio feed. The content loaded onto the set top terminal is transcoded using a transcoding application. If the client does not have an appropriate transcoding application, such transcoding application is downloaded from a server within the system. The transcoding application is used to transcode the loaded content into a desired player format. For example, if the loaded content comprises streaming video and audio at baseband, then such streaming video and audio may be encoded according to any one of the formats previously discussed (e.g., AVI, MPEG, etc.). In the case of the locally loaded content being encoded

according to a first format that is not desired, then the content so encoded is transcoded using the transcoding application to produce content in the desired player format. The transcoded local content is then encapsulated in a desired transport format. The desired transport format is a transport format adapted to a particular access network. The particular  
5 formats associated with various access networks are described above. The transport encapsulated content is then encapsulated in a realtime protocol packet, such as described above. The RTP packet stream is then uploaded to the server for subsequent distribution to clients (i.e., set-top terminals or computers) utilizing the desired player format via access networks utilizing the desired access network transport format.

10 In addition to uploading encapsulated data to the server, the client may also provide access right data to the server for subsequent use in determining who may view the content, who may use the content, how long the content may be viewed, how long the content may be used, which geographic region the viewer or user is in, which set or sub-set of clients within a particular system are to have access, which passwords, if any, are to be used and so  
15 on.

As an example, in the case of a set-top terminal associated with a user wishing to share video imagery of a new baby, the video imagery and associated audio information may be input to a set-top terminal from the video camera. The set-top terminal then uses a coding or transcoding application to encode the video and audio information according to a  
20 desired format, such as the above-mentioned AVI format. Assuming that the subscribers within a system who are to receive this imagery (e.g., friends and neighbors) are within a system having an access network utilizing the MPEG-2 transport format, the set-top terminal or client will then transport encode the AVI encoded content to produce an MPEG-2 transport stream. The MPEG-2 transport stream comprising the AVI encoded  
25 content will be further transport encoded according to the realtime protocol techniques described above. The RTP encoded content and associated access information (e.g., passwords for family members and the like) is then uploaded to the server for subsequent distribution on demand to the appropriate viewers or users.

Although various embodiments which incorporate the teachings of the present invention have been shown and described in detail herein, those skilled in the art can readily devise many other varied embodiments that still incorporate these teachings.

What is claimed is:

1. Method for preprocessing content for a stream caching server in an interactive information distribution system, said method comprising:
  - 5 retrieving content in a first subscriber terminal;
  - transcoding said retrieved content into a plurality of MPEG packets;
  - uploading said transcoded content a http server coupled to an access network;
  - encapsulating said transcoded content in accordance to an Internet Protocol (IP) format supported by said stream caching server; and
  - 10 transmitting said encapsulated content for storage in said stream caching server.
2. The method of claim 1 further comprising:
  - downloading an applet to said first subscriber terminal from a http server; and
  - executing said applet to initiate said retrieving, said transcoding and said uploading.
- 15 3. The method of claim 1 further comprising:
  - creating metadata for said content, where said metadata comprises indexing information used by said stream caching server in response to a command provided by a user viewing said content at a second subscriber terminal; and
  - 20 uploading said metadata with said content.
4. The method of claim 3 wherein said metadata is encapsulated with said transcoded content in said IP format.
- 25 5. The method of claim 4 wherein said command comprises at least one of fast forward (FF), rewind (REW), pause, stop, bookmark, and return to place.
6. The method of claim 1 where said retrieved content in said first subscriber terminal is one of an AVI file, a MPEG-1 file and a moving JPEG file.

30

7. The method of claim 1 wherein said plurality of MPEG packets is contained in a payload of an IP packet.

8. The method of claim 7 wherein said plurality of MPEG packets is contained in a payload of a Realtime Transfer Protocol (RTP) packet contained in a payload of an IP packet.

9. The method of claim 1 wherein said plurality of MPEG packets comprises a plurality of one of a MPEG-2 packet and a MPEG-4 packet.

10

10. The method of claim 1 said IP formatted content is retrieved from said streaming cache server in response to a request for content from a second subscriber terminal, and streamed via a distribution network and said access network to said second subscriber terminal.

15

11. The method of claim 1 said IP formatted content is retrieved from said streaming cache server in response to a request for content from another stream cache server, and streamed to from said caching server to that other caching server.

12. The method of claim 1 wherein said retrieving of RTP formatted content and said streaming are conditioned upon a user of said second subscriber terminal providing a correct password to a http server as configured by a user of said first subscriber terminal.

13. The method of claim 1 wherein said access network comprises one of a wide area network, a local area network, a cable network, a carrier network, a satellite network and a wireless terrestrial network.

14. A system for preprocessing content for a stream caching server in an interactive information distribution system, said system comprising:

a first subscriber terminal for receiving content, transcoding said content into a plurality of MPEG packets, and uploading said transcoded content to an access network; and

a digital link for encapsulating said transcoded content in accordance to an Internet  
5 Protocol (IP) supported by said stream caching server, and transmitting said encapsulated content to said stream caching server.

15. The system of claim 14 further comprising:

a http server, coupled to said access network, for providing an applet to said first  
10 terminal and for providing a user interface for a user of said first subscriber terminal.

16. The system of claim 15 wherein said first subscriber terminal downloads an applet from said http server, and executes said applet to initiate said receiving, transcoding and uploading.

15

17. The system of claim 14 wherein said first terminal creates metadata for said content upon executing said applet, where said metadata comprises indexing information used by said caching stream server in response to a command provided by a user viewing said content at a second subscriber terminal.

20

18. The system of claim 17 wherein said metadata is encapsulated with said IP content in said IP format.

19. The system of claim 17 wherein said command comprises at least one of fast  
25 forward (FF), rewind (REW), pause, stop, bookmark, and return to place.

20. The system of claim 14 wherein said plurality of MPEG packets is contained in a payload of an IP packet.



21. The system of claim 20 wherein said plurality of MPEG packets is contained in a payload of a Realtime Transfer Protocol (RTP) packet contained in a payload of an IP packet.

5 22. The system of claim 14 wherein said plurality of MPEG packets comprises a plurality of one of a MPEG-2 packet and a MPEG-4 packet.

23. The system of claim 15 further comprising:  
a second subscriber terminal for sending a request for content to said http server and  
10 for receiving said content retrieved from said stream caching server and streamed via a distribution network and said access network.

24. The system of claim 14 further comprising:  
a remote stream caching server for streaming said content to said stream caching  
15 server in response to content from said stream caching server.

25. The system of claim 14 wherein said retrieval of RTP formatted content and said streaming are conditioned upon a user of said second subscriber terminal providing a correct password to a http sever as configured by a user of said first subscriber terminal.  
20

26. The system of claim 14 wherein said access network comprises one of a wide area network, a local area network, a cable network, a carrier network, a satellite network, and a terrestrial wireless network.

25 27. A method for use in a client server system, comprising:  
loading, into a client, content local to said client;  
loading into said client as necessary, the transcoding application from said server,  
said transcoding application operative to transcode or encode content into a desired player  
format;  
30 transcoding or encoding said loaded content into said desired player format;

encapsulating said transcoded content into a desired transport format; and  
uploading said encapsulated content to said server.

28. The method of claim 27, further comprising the step of:

5 uploading to said server access rights associated with said encapsulated data.

29. The method of claim 28, wherein said access rights comprise at least one of a  
password protection scheme, a time-to-view parameter, a time-to-use parameter, a defined  
use population and a defined geographic population.

10

30. The method of claim 26, wherein said step of encapsulating comprises the steps of:  
first encapsulating said transcoded content according to a transport format adapted  
to a predefined access network; and

further encapsulating said transport formatted content within a realtime protocol  
15 (RTP) packet adapted to an internet protocol (IP) network.

1/6

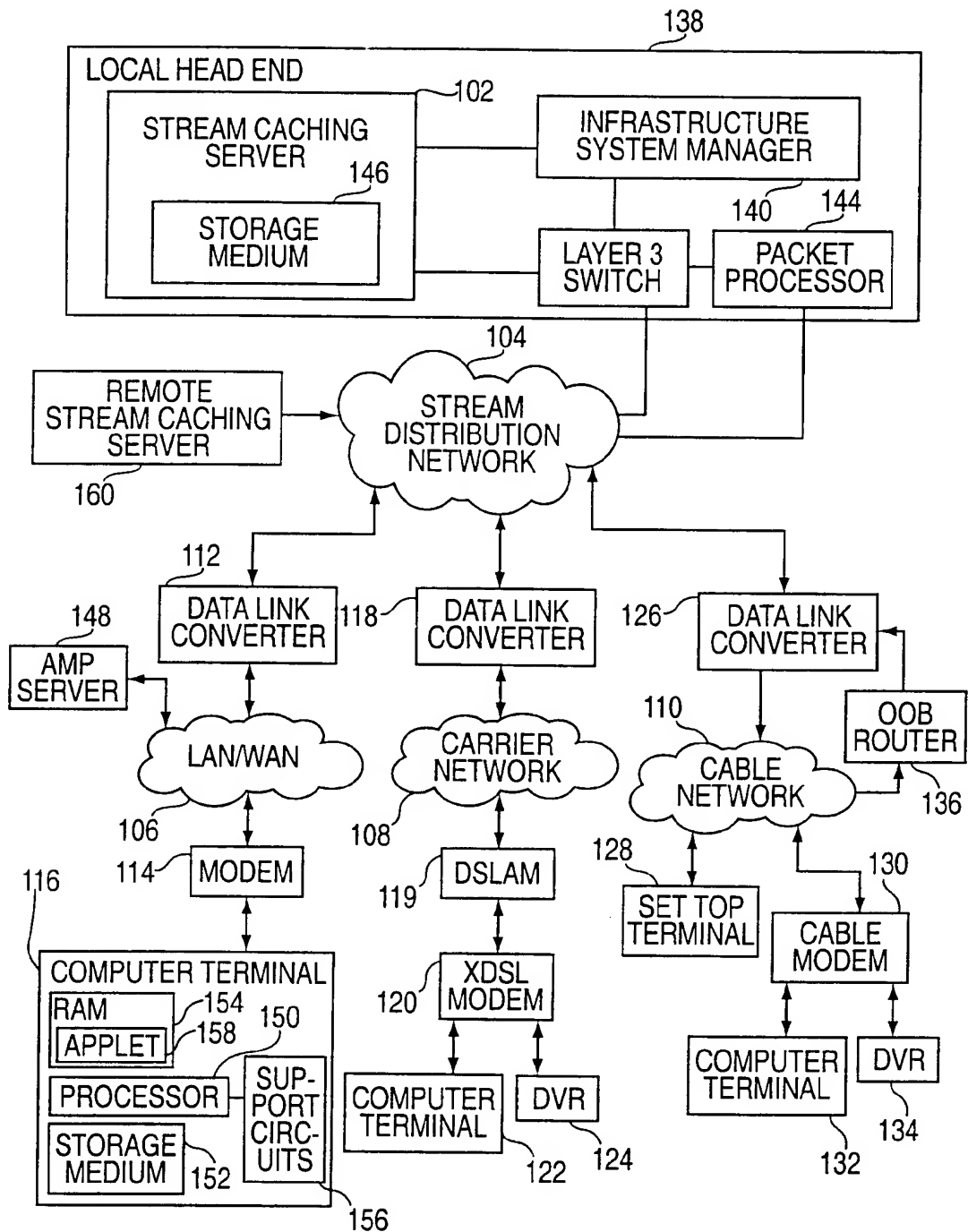


FIG. 1

2/6

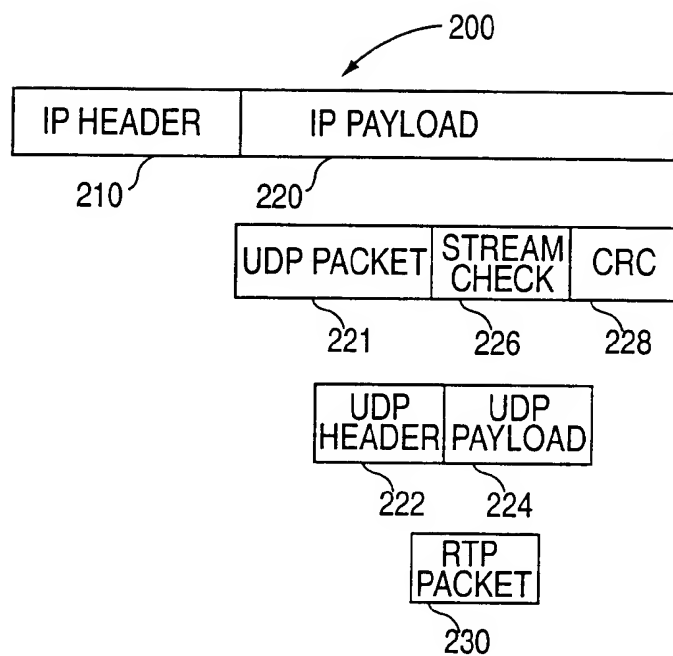


FIG. 2A

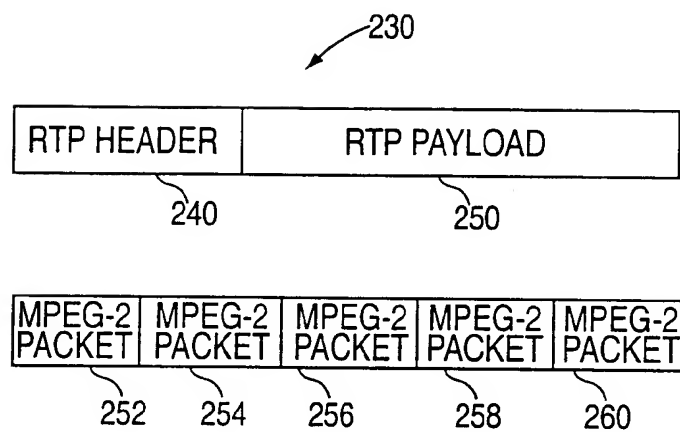


FIG. 2B

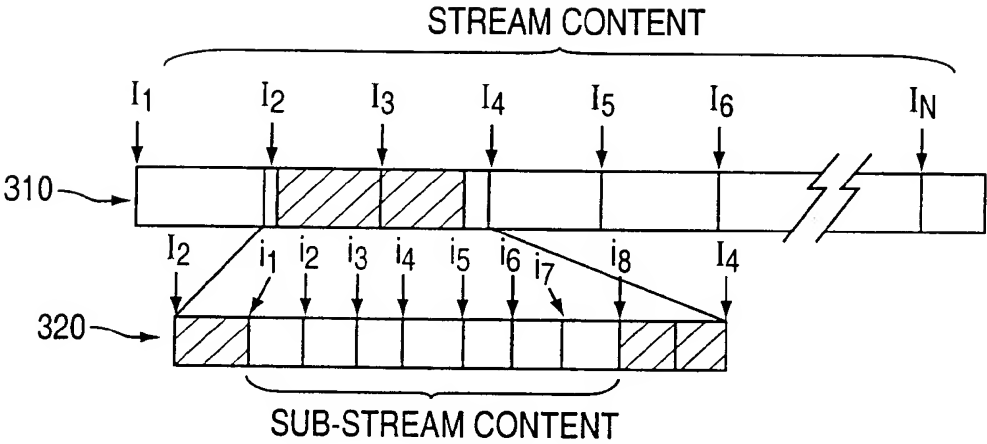


FIG. 3

4/6

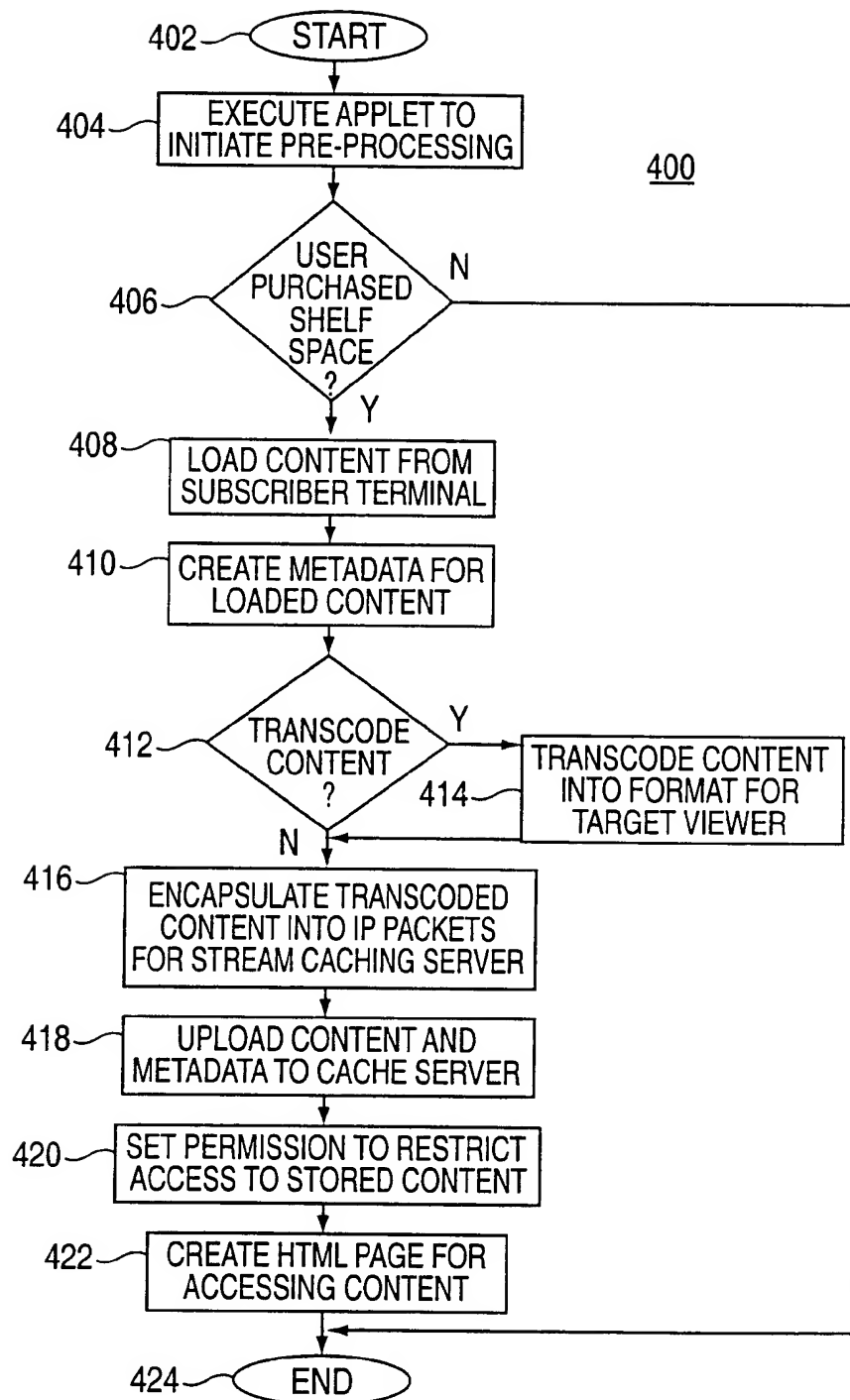


FIG. 4

5/6

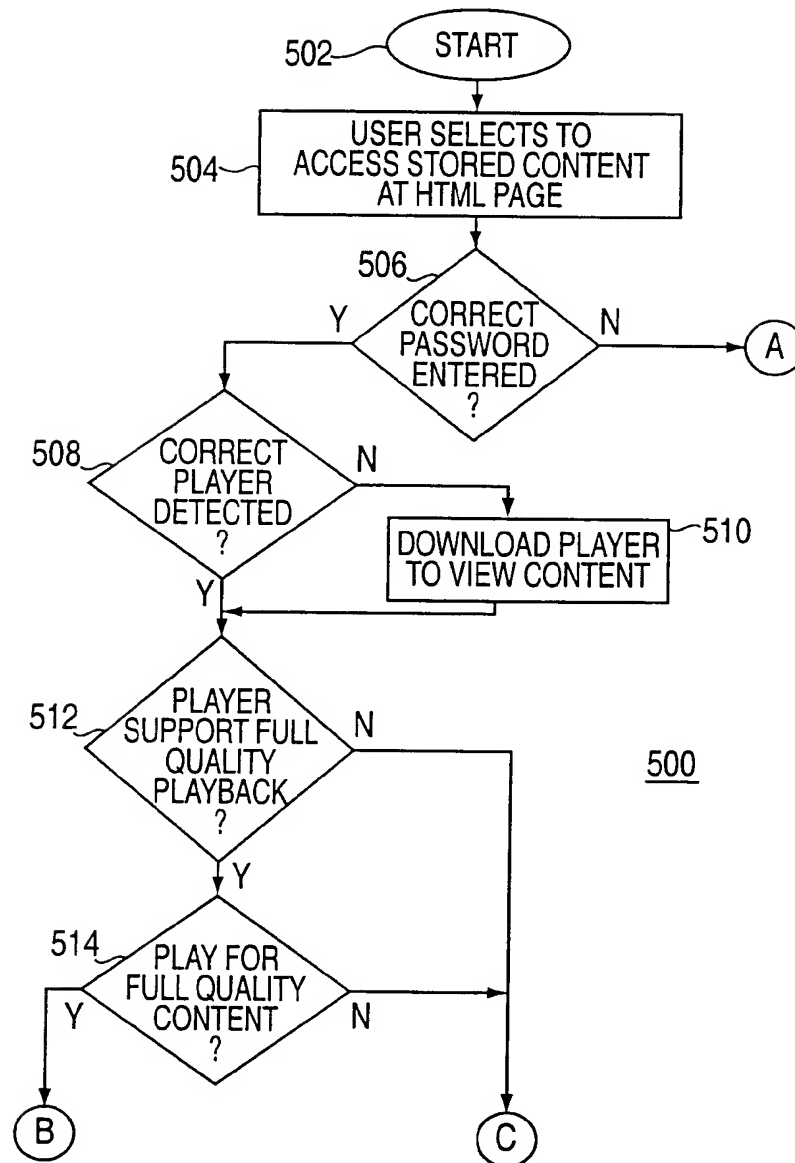


FIG. 5A

6/6

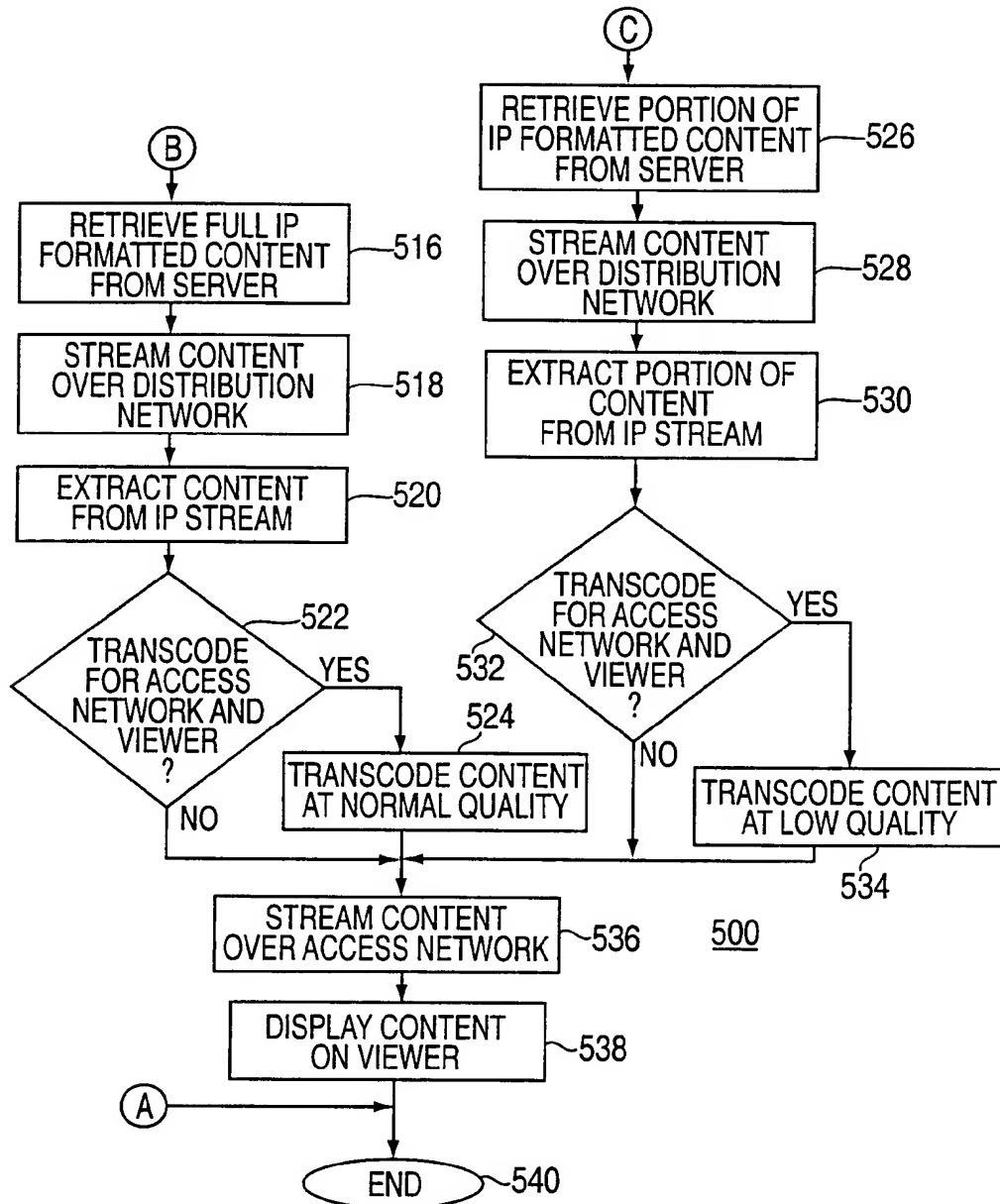


FIG. 5B



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US01/02802

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) : G06F 15/16

US CL : 709/219, 217, 229

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 709/219, 217, 229

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

USPT, PGPB, JPAB, EPAB, DWPI, TDBD

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,838,678 A (DAVIS et al) 17 November 1998, col. 9 line 55 through col. 10 line 8 and figures 8-9.	1-30
Y	US 5,856,973 A (THOMPSON) 05 January 1999, col. 3 line 29 et seq, col. 8 lines 47-67, and columns 9-11.	1-30
Y,P	US 6,079,566 A (ELEFThERIADIS et al) 27 June 2000, col. 2 lines 1-47, and columns 3-7.	1-30



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

19 APRIL 2001

Date of mailing of the international search report

04 MAY 2001

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

MEHMET GECKE *James R. Matthews*

Telephone No. (703) 305-9676